

APPENDIX

31. (Once Amended) A file content classification system comprising:
a plurality of agents, each agent including a file content ID generator creating file content IDs using a mathematical algorithm, at least one agent provided on one of a plurality of clients;
an ID appearance database, provided on a server, coupled to receive file content IDs from the agents; and
a characteristic comparison routine on the server, identifying a characteristic of the file content based on the appearance of the file content ID in the appearance database and transmitting the characteristic to the client agents.
38. (Once Amended) The content classification system of claim [39] 31 wherein said characteristic comprises junk e-mail and said characteristic is defined by a frequency of appearance of a file content ID.
39. (Once Amended) A method for identifying characteristics of data files, comprising:
receiving, on a processing system, file content identifiers for data files from a plurality of file content identifier generator agents, each agent provided on a source system and creating file content IDs using a mathematical algorithm, via a network;
determining, on the processing system, whether each received content identifier matches a characteristic of other identifiers; and
outputting, to at least one of the source systems responsive to a request from said source system, an indication of the characteristic of the data file based on said step of determining.
46. (Once Amended) A method of filtering an email message, comprising:
receiving, on a second computer, a digital content identifier created using a mathematical algorithm unique to the message content from at least two of a plurality of first computers having digital content ID generator agents;

comparing, on the second computer, the digital content identifier to a characteristic database of digital content identifiers received from said plurality of first computers to determine whether the message has a characteristic; and

responding to a query from at least one of said plurality of computers to identify the existence or absence of said characteristic of the message based on said comparing.

Please cancel claims 51 – 54.

56. (Once Amended) A method for providing a service on the Internet, comprising:

collecting data on a processing system from a plurality of systems having a client agent generating digital content identifiers created using a mathematical algorithm for each of a plurality of files on the Internet to a server having a database;

characterizing the files on the server system based on said digital content identifiers received relative to other digital content identifiers collected in the database; and

transmitting a [content] substance identifier from the server to the client agent indicating the presence or absence of a characteristic in the file.